

Overconstrained dynamics in galaxy redshift surveys

Mikel Susperregi

Fisika Teorikoaren eta Zientzi Historiaren Saila, Zientzi Fakultatea, Euskal Herriko Unibertsitatea, PO Box 644, 48080 Bilbao, Spain
 Email: wtpsuxxm@ehu.es

Version of 1 February 2008

ABSTRACT

The least-action principle (LAP) method is used on four galaxy redshift surveys to measure the density parameter Ω_m and the matter and galaxy-galaxy power spectra. The datasets are PSCz, ORS, Mark III and SFI. The LAP method is applied on the surveys simultaneously, resulting in an overconstrained dynamical system that describes the cosmic overdensities and velocity flows. The system is solved by relaxing the constraint that each survey imposes upon the cosmic fields. A least-squares optimization of the errors that arise in the process yields the cosmic fields and the value of Ω_m that is the best fit to the ensemble of datasets. The analysis has been carried out with a high-resolution Gaussian smoothing of 500 km s^{-1} and over a spherical selected volume of radius $9,000 \text{ km s}^{-1}$. We have assigned a weight to each survey, depending on their density of sampling, and this parameter determines their relative influence in limiting the domain of the overall solution. The influence of each survey on the final value of Ω_m , the cosmographical features of the cosmic fields and the power spectra largely depends on the distribution function of the errors in the relaxation of the constraints. We find that PSCz and Mark III are closer to the final solution than ORS and SFI. The likelihood analysis yields $\Omega_m = 0.37 \pm 0.01$ to 1σ level. PSCz and SFI are the closest to this value, whereas ORS and Mark III predict a somewhat lower Ω_m . The model of bias employed is a scale-dependent one, and we retain up to 42 bias coefficients b_{rl} in the spherical harmonics formalism. The predicted power spectra are estimated in the range of wavenumbers $0.02 h \text{ Mpc}^{-1} \lesssim k \lesssim 0.49 h \text{ Mpc}^{-1}$, and we compare these results with measurements recently reported in the literature.

Key words: cosmology: theory – large-scale structure of Universe – cosmic flows – galaxies: distances, velocities and redshifts

1 INTRODUCTION

Galaxy redshift surveys are elements of prime importance in the determination of cosmological parameters. Looking at redshift-space distortions we are able to obtain maximum likelihood estimates of Ω_m (Fisher, Scharf & Lahav 1994; Heavens & Taylor 1995; Baker et al. 1998; Hamilton 1998 and references therein; Dekel 1999a,b; Hamilton, Tegmark & Padmanabhan 2000) and the galaxy-galaxy power spectrum (Fisher et al. 1993; Feldman, Kaiser & Peacock 1994; Lin et al. 1996; Sutherland et al. 1999; Hamilton & Tegmark 2000). The dynamics of cosmic velocity flows is also another way to measure Ω_m (Willick et al. 1997; Dekel, Burstein & White 1997; da Costa et al. 1998; Kashlinsky 1998; Sigad et al. 1998; Susperregi 2001, hereafter S01) and the power spectrum (Strauss & Willick 1995 and references therein; Kolatt & Dekel 1997; Willick et al. 1997; Freudling et al. 1999; Zehavi & Dekel 1999). In the larger picture, galaxy redshift surveys contribute modestly, alongside a myriad of other datasets (CMB, SN1a data, weak lensing surveys, etc) to the

undertaking of estimating the key cosmological parameters (Ω_m , Ω_0 , Ω_B , Ω_Λ or Ω_Q , h , n_s , n_t , etc)(for a review of how this inventory of over a dozen parameters is currently tackled, see e.g. Turner 1999; Primack 2000); all the datasets, including galaxy surveys, can be ideally combined by exploiting cosmic complementarity (Eisenstein, Hu & Tegmark 1998,1999) leading to demonstrably concordant, or at least not inconsistent, predictions.

A number of techniques have been employed to maximize the parameter information extracted from surveys. The Fisher information matrix approach is one example of this (Tegmark 1997; Tegmark, Taylor & Heavens 1997; Goldberg & Strauss 1998; Taylor & Watts 2000). This approach is sound and will lead to very accurate results when applied to forthcoming surveys such as SDSS and 2dF; the computational effort can be successfully minimised in those cases via data-compression techniques, e.g. via the Karnhunen-Loève “signal-to-noise” eigenmodes (Vogeley & Szalay 1996; Tegmark et al. 1998; Matsubara, Szalay & Landy 2000), quadratic compression into high resolution powers (Tegmark

1997; Tegmark, Taylor & Heavens 1998; Tegmark et al. 1998; Padmanabhan, Tegmark & Hamilton 2000), etc. There is, on the other hand, a case to be made for combining different galaxy redshift surveys and extracting information jointly from them, as an ensemble, prior to comparing them with different kinds of datasets. As a step prior to exploiting cosmic complementarity, it can be argued that redshift surveys are best taken into account as an ensemble than individually, however larger and better sampled is any given dataset in comparison to its predecessors. It is to be expected that a comparison of surveys leaves considerable room to manoeuvre to best rid of, or, at least compensate for, the errors inherent in each dataset, stochasticity, unaccounted for contaminants, etc.

The goal of this paper is to set out a procedure to extract an accurate estimate of Ω_m from galaxy redshift surveys, when the datasets are studied as an ensemble. The measurement of Ω_m will then be the optimal estimate for all the surveys considered. The method will also enable us to estimate the matter and galaxy-galaxy power spectra. The backbone of the procedure is essentially the least-action principle (LAP) reconstruction employed in S01. That paper showed that the LAP method is efficient to measure Ω_m from a given galaxy redshift survey and that it conveniently breaks the degeneracy between Ω_m and the bias, thus permitting to measure Ω_m within fairly arbitrary bias schemes. In this paper we consider a bias relationship that is scale-dependent in order to investigate the dynamics more realistically. Bias is in any case stochastic as well as scale-dependent (Pen 1998; Tegmark & Peebles 1998; Dekel & Lahav 1999; Tegmark & Bromley 1999; Taruya 2000), and the bias relationship is fully determined by the distribution $P(\delta_g|\delta_m)$ (Dekel & Lahav 1999; Sigad, Branchini & Dekel 2000). The stochastic ingredient is largely sample-dependent, and the local scatter that is the measure of stochasticity in $P(\delta_g|\delta_m)$ is a characteristic of each survey, its shot noise and environmental effects. The whole point of studying an ensemble of surveys is to minimise this effect, and therefore for simplicity we neglect the existence of stochasticity in the bias and we focus on its scale dependence.

The galaxy redshift surveys that we examine in this paper are PSCz, ORS, Mark III and SFI. The LAP method is initially applied to each survey individually in the manner of S01 with a scale-dependent bias model. Hence one obtains the cosmic fields and a measurement of Ω_m for each survey. Each survey is regarded as a constraint on the cosmic fields, and as they are all indeed different, the system that results from considering all constraints simultaneously is unavoidably overconstrained. Strictly speaking it is also incompatible. The method to find an optimal solution that is a meaningful representation of the ensemble entails relaxing all the constraints. The errors that this relaxation brings in are dealt with the least-square method, and this yields the best fit. The procedure gives us also an estimate of Ω_m , which is really a joint measurement for all the surveys in the ensemble. From the cosmic fields and the bias relationship it is then straightforward to compute the power spectra, higher-order correlation functions, etc.

The article is structured as follows. A brief description of the LAP method is given in §2 and the algebra is relegated to the appendices, though the necessary steps to undertake the numerical implementation are set out in §2.1; the char-

acteristics of the surveys that are relevant to this paper are summarized in §3; the procedure for finding LAP solutions for the ensemble of datasets is described in §4 for the most general case; in §5 we discuss the application of the method on the four surveys in question and their results; finally §6 discusses the main results of this paper, in comparison with similar measurements in the literature.

2 LAP PRELIMINARIES

The three ingredients of the LAP method are: (a) two boundary conditions and (b) the Lagrangian of the self-gravitating matter field. The two boundary conditions largely determine the cosmological model, as they are the endpoints in the evolution of our Universe. Initial homogeneity, as measured by the Sachs-Wolfe constraint on the CMB, suggests the first is

$$\delta(t \rightarrow 0, \mathbf{x}) \approx 0, \quad (1)$$

where δ is the matter density contrast. The statistics of the primordial δ is thus unconstrained. The second boundary condition is the Universe as is represented in the current epoch by a galaxy redshift survey; this is characterised by the galaxy number-count $n(\mathbf{s})$ (\mathbf{s} denotes redshift coordinates, (cz, θ, φ)). This constraint is given by (S01)

$$\rho_s(\mathbf{s}) = x^2 \frac{N_{\text{gals}}}{V} \left[\frac{1 + g_0(\mathbf{x})}{1 + \alpha_0(\mathbf{x})''} \right], \quad (2)$$

where ρ_s relates to the galaxy number-count via $\rho_s \equiv dn(\mathbf{s})/dsd\Omega$ (where $d\Omega$ denotes a unit of solid angle), V is the volume of the survey, x is the radial comoving distance, α is the velocity potential (i.e. $\mathbf{v} \equiv \nabla\alpha$), the prime denotes the line-of-sight derivative d/dx , $g \equiv \delta_{\text{gals}}$ is the galaxy number density contrast and the subindex 0 denotes the present time. We use comoving quantities consistently throughout the paper. The bias model that we shall adopt is scale-dependent, such that

$$g_0(\mathbf{k}) = b(k)\delta_0(\mathbf{k}). \quad (3)$$

The method is used following the same procedure as that expounded in S01, with the addition of a scale-dependent bias. The relevant LAP equations are summarized in Appendix A for reference.

2.1 Numerical resolution

We apply the algorithm as follows.

- The dataset \mathcal{D} is transformed into a z -space field $\rho_s(\mathbf{s})$, by computing the discrete derivatives of the galaxy number-count $n(\mathbf{s})$, binned to the smallest redshift spacing Δz , and then the result is smoothed choosing an appropriate smoothing length r_s . We shall only implement Gaussian smoothing, $W(k) = \exp(-k^2 r_s^2/2)$.

- The goal is to find a best fit for the modes $\delta_y(t)$, $\alpha_y(t)$ (where $y \equiv rlm$ in the spherical harmonics and Bessel functions decomposition, following the formalism of Appendix A). A consistent method to achieve this is to get progressively closer to the real solution by trying Ansätze of increasing non-linearity in successive iterations. The starting point

is the linear solution, which is approximated by inverting the relation $\delta_s \propto -\nabla^2 \alpha_0$, where

$$\delta_s \equiv (s_{\max}/4\pi N_{\text{gals}})\rho_s - 1. \quad (4)$$

For the linear δ field, we estimate

$$\delta_0 \propto \delta_s(\mathbf{x} + \hat{\mathbf{x}}\alpha'_0). \quad (5)$$

In the surveys where the radial velocities are part of the observational input, in our case Mark III and SFI, the v_r data are used to construct the first Ansatz for α'_0 , which is then used to compute (5).

- The linear fields $\delta_{y,0}, \alpha_{y,0}$ are the first Ansatz that we input in the LAP system (A19). We transform these first into the coefficients $\delta_y^{(n)}, \alpha_y^{(n)}$, using (A12), (A13). These are the coefficients of the polynomials of order N that approximate the fields. This initial Ansatz is equivalent to linear theory and it permits us to solve the homogeneous system. It requires an initial guess of the value of Ω_m , as the matrix coefficients S^δ depend on it. It is worthwhile to experiment with a ample range of values such as $0.1 \lesssim \Omega_m \lesssim 1.0$, to examine the impact of these variations in the convergence of the solutions. The solution obtained from the homogeneous system is least-square fitted to (1),(2). This in turn requires an bias model, and from (3) using Parseval's theorem (i.e. $\langle f \cdot g \rangle = \langle f \rangle \cdot \langle g \rangle$, where $\langle f \rangle$ denotes the Fourier transform of f), we have that $g(\mathbf{x}) = b(x)\delta(\mathbf{x})$. In the spherical harmonics and Bessel function decomposition this means

$$g(t, \mathbf{x}) = \sum_y b_{rl}\delta_y(t)j_l(k_r x)Y_{lm}(\theta, \varphi). \quad (6)$$

This assumes that the effect of bias reflects on a superposition of radial modes that are rotationally invariant. In this paper we have fitted b_{rl} for $r \leq 5$ and $l \leq 6$. Our starting point Ansatz in (2) is $b_{rl} = 1.0$.

- The resulting δ_y, α_y are then used to construct the RHS of (A19), the inhomogeneous system. These quadratic terms represent the coupling among the normal modes of the perturbations, and their amplitude is a measure of the degree of non-linearity present.

- Successive iterations thus solve (A19) and we combine this operation with the least-square fitting the solution to the constraints (1),(2). The procedure eventually yields the correct δ_y, α_y if the errors in satisfying (1), (2) are monotonously decreasing after each iteration. We use variations around the solution to construct the Fisher information matrix

$$\mathcal{F}_{ij} \equiv \left\langle \frac{\partial^2 \mathcal{L}(\delta, \alpha)}{\partial \theta_i \partial \theta_j} \right\rangle = - \int d\rho_s \frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \rho_s(\delta, \alpha), \quad (7)$$

where $\theta_0 \equiv \Omega_m$ and $\theta_i \equiv b_{rl}$ for $i > 0$, conveniently ordering the b_{rl} in a sequence. The likelihood function L , is defined as $L \equiv \exp(-\mathcal{L})$. (7) enables us to compute the relative likelihood of estimates of the parameters, particularly Ω_m . The working assumption is that δ_s is in a random field, resulting from the galaxy number-count, which is modelled as a 3D Poisson distribution. This enables us to compute \mathcal{L} . The strength of the LAP method is that it does not require a prior, such as the CDM power spectrum, to estimate the likelihood of the parameters.

3 DATASETS

3.1 PSCz

The sample contains 14,819 galaxies, within a spherical region of radius $x_{\max} \gtrsim 45,000 \text{ km s}^{-1}$ (its median is at $\approx 8,500 \text{ km s}^{-1}$ and the galaxy number-count tails off sparsely up to $\gtrsim 30,000 \text{ km s}^{-1}$), extracted from the IRAS Point Source Catalogue (Saunders et al. 2000), following the selection procedure given by conditions $\log(f_{60}/f_{25}) > 0.5$ and $\log(f_{100}/f_{60}) < 0.75$ (where f_α denotes the measured flux at α microns). The former discards all stars whilst the latter excludes galactic cirrus located away from the Zone of Avoidance (ZoA). The result retains most galaxies plus about 300 mixed contaminants, which are excluded on a one-to-one basis. The source catalogue contains 17,060 sources, of which 1,593 are discarded as they are either within the Milky Way or repeated entries; a further 648 are rejected as very faint galaxies and unidentifiable sources. The resulting 6,500 are a combination of IRAS 1.2 Jy and QDOT galaxies plus 3,000 additional sources. The sky coverage is $\sim 84\%$ (the ZoA spans a non-axially symmetric region of $\langle |b| \rangle \lesssim 5^\circ$). We adopt the unmasked survey, therefore maximising the sky coverage and a selected spherical subvolume of radius $x_{\max} \sim 15,000 \text{ km s}^{-1}$. General descriptive information and access to data are available at the PSCz webpage <http://www-astro.physics.ox.ac.uk/~wjs/pscz.html>.

3.2 ORS

The sample contains 8,457 optically selected galaxies, of an angular size $\lesssim 1.9$ arcmin, located within an approximately spherical region $\sim 18,000 \text{ km s}^{-1}$ (Santiago et al. 1995; Hudson et al. 1995; Baker et al. 1998). The dataset is densely sampled within $x_{\max} \sim 8,000 \text{ km s}^{-1}$, which is the spherical subvolume that we shall consider, and it is sparse in the region $x \gtrsim 8,000 \text{ km s}^{-1}$. The sample is a compilation of optically-selected galaxies extracted from the Upsala General Catalogue (UGC; $\delta \geq -2.5^\circ$), European Southern Observatory (ESO; $\delta < -17.5^\circ$) catalogue and Extension to the Southern Observatory Catalogue (ESGC; covering the remaining region just south of the Celestial Equator. The sky coverage of these samples is respectively 34%, 19% and 9%. ORS shares a number of galaxies in common with PSCz though it contains a larger fraction of spirals. The predicted velocity fields in ORS and PSCz are also in good agreement (Baker et al. 1998). The sky coverage is 62%, i.e. significantly more limited than in PSCz (and even IRAS 1.2 Jy), due to the extinction in the ZoA, and it spans over $|b| \gtrsim 20^\circ$. In the sample utilized ORS is filled with 823 IRAS galaxies in order to achieve $|b| \gtrsim 5^\circ$ coverage. A drawback of ORS is its poor uniformity, as extinction affects optically-selected galaxies more than it does IRAS galaxies. Further descriptive information and access to data are available at the URL: <http://www.astro.princeton.edu/~strauss/ors/>.

3.3 Mark III

The sample contains radial velocities of $\gtrsim 3,400$ galaxies (Willick et al. 1995, 1996, 1997), compiled from several sets of S0 spirals and ellipticals. The sky coverage is the entire sky with the exception of the ZoA, at $|b| \gtrsim 20^\circ - 30^\circ$. The data

are located within a spherical region of radius $x_{\max} \sim 6,000$ km s⁻¹ that is well sampled, though it is fairly anisotropic, as it spans to $x_{\max} \sim 8,000$ km s⁻¹ in some directions and is consigned to $x_{\min} \sim 4,000$ km s⁻¹ in others. We consider the spherical subvolume of radius $x_{\max} 8,000$ km s⁻¹. Distances are inferred via Tully-Fisher and $D_n - \sigma$ distance indicators and they entail an error 17–21%. The data must be carefully prepared to correct for Malmquist biases (following the procedure set out in Sigad et al. (1998) and used by Zaroubi, Hoffman & Dekel (1999); also in a improved version by Dekel et al. (1999)). This results in the correction of the distances of 1,200 objects. Further descriptive information and access to data are available at the URL: <http://redshift.stanford.edu/MarkIII/>.

3.4 SFI

The sample contains radial velocities of $\sim 1,300$ Sbc-Sc galaxies, with inclination $\gtrsim 45^\circ$ north of $\delta < -45^\circ$ and galactic latitude $|b| \gtrsim 10^\circ$ (da Costa et al. 1996, 1998; Haynes et al. 1999a,b). The data are distributed within an approximately spherical region of radius $x_{\max} \sim 7,000$ km s⁻¹. Individual distances are computed in a similar manner as in Mark III, and the distance errors are 15–20%. A comparison with Mark III shows a good agreement (da Costa et al. 1996; ditto for the POTENT reconstruction analysis of Dekel et al. 1999) and it is also in agreement with the predicted velocity field of IRAS 1.2 Jy (da Costa et al. 1998).

4 METHOD

The LAP method applied to one survey is completely determined by the following two constraints:

$$\chi_0(\delta) = 0, \quad (8)$$

$$\chi_{\text{survey}}(\delta, \alpha) = 0, \quad (9)$$

where (8) is the condition of homogeneity (1), and (9) contains the data given by the survey (2). Any additional dataset adds one constraint of the type (9). Let us suppose we have N such datasets. Obviously in practice all samples contain systematic and random errors and not all $N+1$ constraints are satisfied exactly by the same solution; therefore, the system $\chi_i = 0$ with $i = 0, 1, \dots, N$ need not have a consistent solution δ, α . To find a solution, the overconstrained system must be relaxed so that the solution satisfies the constraints (9) within the least margin of error, $\chi_i \approx \epsilon_i$, ($i > 0$), and we optimize this by computing the stationary quantity

$$\tilde{\delta} \left[\sum_{i=1}^N w_i \chi_i^2(\delta, \alpha) \right] = 0, \quad (10)$$

where $\tilde{\delta}$ denotes a variation, not a density contrast, and w_i is the weight given to each survey in the sum. From the central limit theorem, we assume that ϵ_i are normally distributed about zero, i.e. $\langle \epsilon_i \rangle = 0$ and $E(\epsilon_i^2) = \sigma_i$. We will therefore adopt the maximum likelihood values $w_i = \sigma_i^{-2}$, where σ_i^2 is the variance of the distribution $P[\chi_i]$. The likelihood function is then

$$L = (2\pi)^{-N/2} \left(\prod_{i=1}^N \sigma_i^{-1} \right) \exp \left(-\frac{1}{2} \sum_{i=1}^N \frac{\chi_i^2}{\sigma_i^2} \right). \quad (11)$$

Table 1. Sampling factor in the surveys

| | x_{\max} | N_{gals} | η |
|----------|--------------------|--------------------|--------|
| PSCz | 15.0×10^3 | 14.8×10^3 | 0.11 |
| ORS | 8.0×10^3 | 8.3×10^3 | 0.41 |
| Mark III | 6.0×10^3 | 3.4×10^3 | 0.39 |
| SFI | 7.0×10^3 | 1.3×10^3 | 0.09 |

x_{\max} is the radius of the selected spherical volume in each survey in units of km s⁻¹, N_{gals} is the number of galaxies and η is defined by (12).

The solution that satisfies (10) and the LAP equations (A19) is the best fit to the N surveys that one can find.

4.1 Domain of application

The first consideration is to determine the radius of the spherical volume to which the LAP solutions will be consigned. Ideally, one would wish to use the largest selected subvolume possible, in this case that of PSCz, $x_{\max} \sim 15,000$ km s⁻¹. This would nonetheless give too little weight to the other datasets, as the constraints $\chi_i \approx 0$ given by the smaller surveys would have little impact in the outcome of δ, α (the fractions of the selected volumes of ORS, Mark III and SFI, with respect to PSCz are, respectively, 0.15, 0.06 and 0.10). On the other hand, x_{\max} should not be too small either, as a large part of data from the larger samples may thus be left unused. An criterion to determine x_{\max}^{LAP} is the following. We assume that the weight each dataset ought to have in determining the volume of the LAP region is given by the density of its sampling, given by the ratio

$$\eta_i \equiv \frac{N_{\text{gals}}^i}{V_i} \left(\sum_{j=1}^N \frac{N_{\text{gals}}^j}{V_j} \right)^{-1}. \quad (12)$$

This measure of sampling is given in Table 1 for the surveys of our study. It shows that ORS and Mark III are the most densely sampled datasets and SFI is the sparsest. PSCz is also on the whole fairly sparsely sampled. The volume of the LAP region is hence the weighed sum

$$V_{\text{LAP}} = \sum_{i=1}^N \eta_i V_i. \quad (13)$$

Using the values given in Table 1, for the case at hand this results in 8,860 km s⁻¹. In fact this seems to be a good figure also for PSCz as the median redshift of that catalogue is at 8,500 km s⁻¹. We shall adopt the round figure $x_{\max}^{\text{LAP}} = 9,000$ km s⁻¹.

4.2 Solutions

We proceed as follows. First, we apply the LAP method on each sample separately on their own domain of application, following the numerical procedure described in §2.1. Hence, for each survey we obtain δ_i, α_i within a spherical volume of radius x_{\max}^i (using the values given in Table 1). This also yields an estimate of Ω_m and an evaluation of the bias coefficients b_{rl} . The results of the measurement of Ω_m in each survey are given in the next section.

Let us consider these solutions, δ_i, α_i , to construct an

Table 2. Likelihood of Ω_m

| Ω_m | P_{PSCz} | P_{ORS} | P_{MarkIII} | P_{SFI} |
|------------|-------------------|------------------|----------------------|------------------|
| 0.20 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.22 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0.24 | 0.00 | 0.02 | 0.00 | 0.00 |
| 0.26 | 0.00 | 0.31 | 0.02 | 0.00 |
| 0.28 | 0.05 | 0.62 | 0.39 | 0.00 |
| 0.30 | 0.42 | 0.89 | 0.70 | 0.08 |
| 0.32 | 0.73 | 1.00 | 0.91 | 0.38 |
| 0.34 | 0.92 | 0.93 | 1.00 | 0.71 |
| 0.36 | 1.00 | 0.74 | 0.93 | 0.94 |
| 0.38 | 0.91 | 0.33 | 0.72 | 1.00 |
| 0.40 | 0.70 | 0.06 | 0.40 | 0.91 |

The value of the likelihood distribution is normalized to unity at its maximum in each survey. We have binned the range $\Omega_m = 0.20 - 0.40$ in $\Delta\Omega_m = 0.02$ intervals and computed the solutions and the likelihood at each step.

initial Ansatz of the overall solution that satisfies (10). This Ansatz is given by

$$\delta = \sum_{i=1}^N \eta_i \delta_i, \quad (14)$$

$$\alpha = \sum_{i=1}^N \eta_i \alpha_i. \quad (15)$$

Thus we adopt the sampling factor as the best weight to average the fields. A few trial and error experiments show this is a good starting point. In order to consign this Ansatz to the domain (13), if $x_{\text{max}}^{\text{LAP}} < x_{\text{max}}^i$ then one discards from the survey i all the data that lie within the shell $x_{\text{max}}^{\text{LAP}} < x < x_{\text{max}}^i$; and if $x_{\text{max}}^{\text{LAP}} > x_{\text{max}}^i$ then one makes the fields identically equal to zero within the range $x_{\text{max}}^i < x < x_{\text{max}}^{\text{LAP}}$, effectively “padding” that shell to match the LAP volume. In the example of this paper, PSCz data are discarded in the range $9,000 \text{ km s}^{-1} < x < 15,000 \text{ km s}^{-1}$, and ORS, Mark III and SFI are padded with zeros within the shell $x_{\text{max}}^i < x < 9,000 \text{ km s}^{-1}$.

The Ansatz (14),(15) is our starting point to compute the LAP solution over the domain (13). The numerical problem is then to solve (A19), introducing variations in the fields δ , α and in the free parameters Ω_m and b_{rl} to eventually pin down the stationary solution in (10). The numerical resolution is carried out following the procedure set out in §2.1, where all references to the constraints (1), (2) are substituted by (10).

5 RESULTS

We apply the method described in §4 to the four datasets in question, PSCz, ORS, Mark III and SFI. For each survey we construct the galaxy number-count density field $\rho_s(\mathbf{s})$ using a high-resolution Gaussian smoothing length of 500 km s^{-1} . First of all, the LAP equations are solved for each survey following §2.1. This results in an estimate of the cosmic fields δ , α . We carry out a likelihood analysis to compute the value of Ω_m in each case, and the same analysis also yields an estimate of the parameters b_{rl} . In order to do so, the underlying assumption is that the errors in the measurements of

Table 3. Bias parameters b_{rl}

| | $r = 0$ | $r = 1$ | $r = 2$ | $r = 3$ | $r = 4$ | $r = 5$ |
|---------|---------|---------|---------|---------|---------|---------|
| $l = 0$ | 0.98 | 1.01 | 1.05 | 1.09 | 1.12 | 1.13 |
| $l = 1$ | 1.02 | 1.08 | 1.11 | 1.12 | 1.15 | 1.18 |
| $l = 2$ | 1.09 | 1.11 | 1.14 | 1.17 | 1.20 | 1.23 |
| $l = 3$ | 1.16 | 1.18 | 1.20 | 1.23 | 1.26 | 1.30 |
| $l = 4$ | 1.22 | 1.26 | 1.29 | 1.31 | 1.33 | 1.36 |
| $l = 5$ | 1.30 | 1.32 | 1.35 | 1.39 | 1.41 | 1.44 |
| $l = 6$ | 1.38 | 1.40 | 1.42 | 1.45 | 1.46 | 1.52 |

Estimated values of b_{rl} in the range $0 \leq r \leq 5$, $0 \leq l \leq 6$. The 1σ errors in these measurements are ≈ 0.01 (in the worst case, that of b_{43} , it is 1.2×10^{-2} , so this is the upper bound in the error bars of these estimates).

the fields and the parameters, $\Delta\delta_y^{(n)}$, $\Delta\alpha_y^{(n)}$, $\Delta\Omega_m$ and Δb_{rl} , follow a multivariate Gaussian distribution. However, we do not assume that the fields themselves are Gaussian (not at the present time, though we find that they are Gaussian at the initial time, to great accuracy, and therefore it is safe to assume Gaussian initial conditions in the likelihood analysis though our formalism does not require us to do so). The results of $P(\Omega_m)$ for each survey are given in Table 2. The maximum likelihood estimates for Ω_m are $\Omega_m^{\text{PSCz}} = 0.36$, $\Omega_m^{\text{ORS}} = 0.32$, $\Omega_m^{\text{MarkIII}} = 0.34$, $\Omega_m^{\text{SFI}} = 0.38$. The 1σ errors in these measurements are of the order $\approx \mathcal{F}_{00}^{-1/2}$, which results in 2.1×10^{-2} , 1.7×10^{-2} , 1.3×10^{-2} and 1.9×10^{-2} , respectively.

5.1 Measurement of Ω_m

The results obtained above are used to construct an initial Ansatz (14), (15) over the domain (13), that we use to solve (A19) subject to (10). We follow the method set out in §4 and get the LAP solution that is the best fit to the four surveys. The likelihood analysis of the solution yields an estimate $\Omega_m = 0.37 \pm 0.01$ within the 1σ level. This result is quite surprising in view of Table 2, due to the fact that the two surveys with the largest η , ORS and Mark III, seem to pin down Ω_m at a lower value, in the range $0.32 - 0.34$, and yet PSCz and SFI predict a Ω_m that is closer to the final value. In this case, the optimization of the morphological features of the overdensity and velocity fields in the LAP solution with those of the four surveys is such that it favours a larger Ω_m . Indeed the surveys PSCz and Mark III are closer to the final solution than ORS and SFI. The latter two yield the highest χ_i^2 in (10), though SFI is in better agreement with the LAP solution than ORS and therefore χ_{SFI}^2 is smaller. PSCz and Mark III have the smaller variance of the four, and therefore this causes that they have a greater weight in determining the cosmography and the cosmological parameters in the solution.

The likelihood analysis for the bias parameters yields the estimates shown in Table 3. We have computed 42 coefficients, truncating the indices at $r = 5$ and $l = 6$. The higher l enables us to probe into the nonlinear scales, though a stringent limit is set by the smoothing scale of 500 km s^{-1} . The high l coefficients are greater than those in the linear range (low l), showing greater bias at smaller scales, a trend that is also present in all four surveys when the LAP method is applied to each one individually.

5.2 Matter power spectrum

Hereafter we use δ^{LAP} and α^{LAP} evaluated at the maximum likelihood values of the parameters, i.e. $\Omega_m = 0.37$ and b_{rl} given in Table 3. We compute the Fourier modes $\delta(\mathbf{k})$ of δ^{LAP} . The matter power spectrum $P(k)$ is then given directly by

$$\langle \delta(\mathbf{k}) \delta^*(\mathbf{k}') \rangle = (2\pi)^3 P(k) \delta(\mathbf{k} - \mathbf{k}'), \quad (16)$$

assuming that the modes $\delta(\mathbf{k})$ are isotropic and homogeneous, and δ on the RHS is a Dirac delta function. Strictly speaking the fields ought be smooth, though we convolute the solutions with $W(k)$, in order to discard residual noise generated in the derivation of the LAP solution. The results for $P(k)$ are shown in Table 4. The galaxy-galaxy power spectrum is given by

$$P_{gg}(k) \equiv b(k)^2 P(k), \quad (17)$$

and the corresponding results are shown in Table 5. We have consigned the analysis to the range of wavenumbers $0.02 h \text{ Mpc}^{-1} \lesssim k \lesssim 0.49 h \text{ Mpc}^{-1}$ that we have binned in 24 equidistant segments in the log scale. The non-linear part of the power spectrum is located at $k \gtrsim 0.3 h \text{ Mpc}^{-1}$, and it is in this range that we expect the LAP method to yield significant information. The resulting velocity dispersion is estimated $\sigma_v \approx 286 \pm 32 \text{ km s}^{-1}$ and $\sigma_8^2 = 0.64 \pm 0.03$. A COBE-normalized flat Λ CDM model, untilted and with $\Omega_m = 0.37$ and $\Lambda = 0.63$, would fit the measured P_{gg} for $\Gamma = 0.21 \pm 0.03$ within the 1σ error margin. The errors in Tables 4 and 5 are 1σ . In the case of Table 5, they are larger as they result from the product of the errors in $P(k)$ and $b(k)$. In addition to the 1σ errors that are the result of the likelihood analysis, another source of error is due to the limited number of δ_y modes, which is propagated in the coordinate transformations $\delta_y \rightarrow \delta_{\mathbf{k}}$ and $b_{rl} \rightarrow b(k)$. These amount to $\lesssim 10\%$ of the total error, and have been accounted for.

Other power spectra of interest are (see e.g. Pen 1998; Tegmark & Peebles 1998; Dekel & Lahav 1999) the galaxy velocity power

$$P_{gv}(k) \equiv r(k)b(k)fP(k) \quad (18)$$

and the velocity-velocity power

$$P_{vv}(k) \equiv f^2 P(k) \quad (19)$$

where $|r(k)| \geq 1$ is a galaxy-velocity correlation coefficient, $f \approx \Omega_m^{0.6}$ is the linear growth rate. Typically $r \approx 1$ (Hamilton, Tegmark & Padmanabhan 2000), though it may depart from unity in the nonlinear regime. Using the measurement of Ω_m given in §5.1 and Table 4, it is straightforward to compute (19). The resulting errors result from the product of $\Delta P(k)$ and $\Delta \Omega_m = 0.01$.

6 DISCUSSION

The measurement of Ω_m that we have obtained by applying the LAP method simultaneously on PSCz, ORS, Mark III and SFI is $\Omega_m = 0.37 \pm 0.01$. This is consistent with the estimates that we have obtained from each survey separately (see Table 2), though ORS and Mark III predict a somewhat lower value, $\Omega_m \approx 0.32 - 0.34$. The main factors in the determination of Ω_m in our analysis are chiefly two:

(1) the variance of the distribution $P[\chi_i]$ defined in §4, (2) the errors χ_i^2 , that reflect the resemblance of cosmographical features in the survey and the LAP solution. The ideal combination is a small variance and a minimal value of χ_i^2 , indicating a close resemblance between the survey and the LAP solution. The surveys that satisfy these requirements best are PSCz and Mark III. The variance of ORS and SFI is not much greater than that of PSCz and Mark III, so these surveys are consistent with the final LAP solution by an error of amplitude no greater than 6%.

These results indicate a higher value of Ω_m than reported in S01 for IRAS 1.2 Jy ($\Omega_m = 0.30$ with a linear bias of $b = 1.1$). In order to compare the results of this paper with other measurements reported in the literature, most of which are in the form of the redshift distortions parameter $\beta \approx \Omega_m^{0.6}/b$, we need to have some information about the optimal value of the linear bias in those measurements. Hamilton (1998) surveys measurements up until mid-1997, most of which fall within the region $\beta \approx 0.45 - 0.75$. A higher measurement of $\beta = 0.89 \pm 0.12$ is reported by Sigad et al. (1998) from Mark III, and recently a much lower one of $\beta = 0.41_{-0.12}^{+0.13}$ by Hamilton, Tegmark & Padmanabhan (2000) from PSCz and a moderately low one of $\beta = 0.5 \pm 0.1$ by Nusser et al. (2000) from a comparison between ENEAR peculiar velocities and the PSCz gravity field. Our results are consistent with β measurements at the lower part of the range, as these allow a density parameter in the region $\sim 0.3 - 0.4$ for reasonable values of the linear bias $b \approx 1.0 - 1.3$. Higher values of β require a much larger Ω_m that is not consistent with our results. Such is the case of Zehavi & Dekel (1999), who predict $\Omega_m \gtrsim 0.6$ for both Mark III and SFI, and a value of $\Omega_m \approx 0.4$ that would be consistent with our results has a low likelihood in their analysis.

In Table 4 and 5 we have given the values of the matter and galaxy-galaxy power spectra respectively. The values of P_{gg} are consistent, within the 1σ errors given, with the results of Sutherland et al. (1999) and Hamilton & Tegmark (2000) for PSCz. The range of wavenumbers considered by Hamilton & Tegmark (2000) is considerable, $0.01 h \text{ Mpc}^{-1} \lesssim k \lesssim 300 h \text{ Mpc}^{-1}$, and they carry out a careful analysis of errors by computing the prewhitened power spectrum. Using similar results, Hamilton, Tegmark & Padmanabhan (2000) fit their measured P_{gg} to a COBE-normalized, untilted flat Λ CDM, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ (Eisenstein & Hu (1998)), within a range of wavenumbers $0.01 h \text{ Mpc}^{-1} \lesssim k \lesssim 1.00 h \text{ Mpc}^{-1}$. Our results is in agreement with theirs, within the more limited range of wavenumbers that we have estimated, namely $0.01 h \text{ Mpc}^{-1} \lesssim k \lesssim 0.49 h \text{ Mpc}^{-1}$, and our analysis favours a slightly higher value of Ω_m that, within the errors, is fairly close to $\Omega_m \approx 0.4$. Furthermore, our analysis predicts a value of the velocity dispersion $\sigma_v = 286 \pm 32 \text{ km s}^{-1}$, slightly lower than one of the values adopted by Sutherland et al. (1999), $\sigma_v = 300 \text{ km s}^{-1}$. The analysis of peculiar velocities in Mark III and SFI, by Zaroubi et al. (1997) and Freudling et al. (1999) respectively (see also Zehavi & Dekel (1999)), yields a high-amplitude estimate of the power spectrum, $P(k = 0.1 h \text{ Mpc}^{-1}) \Omega_m^{1.2} = (4.5 \pm 2.0) \times 10^3 (h^{-1} \text{ Mpc})^3$. Our estimates are consistent with this (from Table 4, we have $P(k = 0.1 h \text{ Mpc}^{-1}) \approx (4.63 \pm 0.36) \times 10^3 (h^{-1} \text{ Mpc})^3$), pinning down the values to greater accuracy. It is also the

Table 4. Matter power spectrum

| k ($h \text{ Mpc}^{-1}$) | $P(k)$ ($h^{-3} \text{ Mpc}^3$) | $\Delta P(k)$ | k ($h \text{ Mpc}^{-1}$) | $P(k)$ ($h^{-3} \text{ Mpc}^3$) | $\Delta P(k)$ |
|---------------------------------|--------------------------------------|---------------|---------------------------------|--------------------------------------|---------------|
| 0.0182 | 9532 | 2266 | 0.1010 | 4630 | 363 |
| 0.0210 | 10043 | 3451 | 0.1165 | 3981 | 298 |
| 0.0242 | 11283 | 4016 | 0.1343 | 3374 | 274 |
| 0.0279 | 12499 | 3220 | 0.1550 | 2935 | 260 |
| 0.0322 | 13274 | 3061 | 0.1787 | 1879 | 184 |
| 0.0372 | 13883 | 2874 | 0.2062 | 1140 | 165 |
| 0.0429 | 14392 | 2249 | 0.2378 | 826 | 149 |
| 0.0494 | 12529 | 1892 | 0.2743 | 576 | 121 |
| 0.0570 | 10302 | 1220 | 0.3164 | 488 | 82 |
| 0.0658 | 8821 | 921 | 0.3649 | 321 | 71 |
| 0.0759 | 8427 | 887 | 0.4209 | 267 | 55 |
| 0.0875 | 8033 | 408 | 0.4855 | 210 | 41 |

Estimated values of $P(k)$ in the interval of wavenumbers $0.0182 h \text{ Mpc}^{-1} \leq k \leq 0.4855 h \text{ Mpc}^{-1}$. The range has been binned to produce 24 measurements within this interval. The errors, $\Delta P(k)$, are to 1σ level in the likelihood estimates.

Table 5. Real-space power spectrum

| k ($h \text{ Mpc}^{-1}$) | $P_{\text{gg}}(k)$ ($h^{-3} \text{ Mpc}^3$) | $\Delta P(k)$ | k ($h \text{ Mpc}^{-1}$) | $P_{\text{gg}}(k)$ ($h^{-3} \text{ Mpc}^3$) | $\Delta P(k)$ |
|---------------------------------|--|---------------|---------------------------------|--|---------------|
| 0.0182 | 4232 | 2389 | 0.1010 | 5013 | 408 |
| 0.0210 | 9721 | 3672 | 0.1165 | 4521 | 353 |
| 0.0242 | 11396 | 4229 | 0.1343 | 3882 | 299 |
| 0.0279 | 13102 | 3309 | 0.1550 | 3621 | 275 |
| 0.0322 | 13855 | 3088 | 0.1787 | 2288 | 192 |
| 0.0372 | 14935 | 2946 | 0.2062 | 1720 | 174 |
| 0.0429 | 16223 | 2321 | 0.2378 | 1211 | 165 |
| 0.0494 | 13725 | 1922 | 0.2743 | 1088 | 132 |
| 0.0570 | 11263 | 1321 | 0.3164 | 975 | 89 |
| 0.0658 | 8860 | 974 | 0.3649 | 720 | 80 |
| 0.0759 | 8751 | 925 | 0.4209 | 612 | 62 |
| 0.0875 | 8231 | 530 | 0.4855 | 488 | 53 |

Estimated values of P_{gg} . The errors are, like in Table 4, 1σ .

case that in our estimates of $P_{\text{gg}}(k)$ the error bars are significantly smaller than in Hamilton & Tegmark (2000) and Sutherland et al. (1999).

ACKNOWLEDGEMENT

This research has been funded by the Eusko Jaurlaritza research fellowship BFI99.116 and in part at EHU by research grant UPV172.310-G02/99.

REFERENCES

- Abramowitz M., Stegun I.A., 1965, *Handbook of Mathematical Functions*, Dover Inc., New York
- Baker J.E., Davis M., Strauss M.A., Lahav O., Santiago B.X., 1998, ApJ 508, 6
- Courant R., Hilbert D., 1989, *Methods of Mathematical Physics*, John Wiley & Sons, New York (1st English edition, 1953)
- da Costa L.N., Freudling W., Wegner G., Giovanelli R., Haynes M.P., Salzer J.J., 1996, ApJ 468, L5
- da Costa L.N., Nusser A., Freudling W., Giovanelli R., Haynes M.P., Salzer J.J., Wegner G., 1998, MNRAS 299, 425
- Dekel A., 1999a, in *Formation of Structure in the Universe*, ed. A. Dekel and J.P. Ostriker, Cambridge University Press, p. 250
- Dekel A., 1999b, in *Cosmic Flows: Towards an Understanding of Large-Scale Structure*, eds. S. Courteau, M.A. Strauss and J.A. Willick, ASP Conf. Series
- Dekel A., Burstein D., White S.D.M., 1997, in *Critical Dialogues in Cosmology*, ed. N.G. Turok, World Scientific, Singapore, p. 175
- Dekel A., Eldar A., Kolatt T., Yahil A., Willick J.A., Faber S.M., Courteau S., Burstein D., 1999, ApJ 522, 1
- Dekel A., Lahav O., 1999, ApJ 520, 24
- Eisenstein D.J., Hu W., 1998, ApJ 496, 605
- Eisenstein D.J., Hu W., Tegmark M., 1998, ApJ 504, L57
- Eisenstein D.J., Hu W., Tegmark M., 1999, ApJ 518, 2
- Feldman H.A., Kaiser N., Peacock J.A., 1994, ApJ 426, 23
- Fisher K.B., Davis M., Strauss M.A., Yahil A., Huchra J.P., 1993, ApJ 402, 42
- Fisher K.B., Scharf C.A., Lahav O., 1994, MNRAS 266, 219
- Freudling W., Zehavi I., da Costa L.N., Dekel A., Eldar A., Giovanelli R., Haynes M.P., Salzer J.J., Wegner G., Zaroubi S., 1999, ApJ 523, 1
- Goldberg D.M., Strauss M.A., 1998, ApJ 495, 29
- Hamilton A.J.S., 1998, in *The Evolving Universe*, Hamilton D.,

- ed., Kluwer, Dordrecht, p. 185; *astro-ph/9708102*
- Hamilton A.J.S., Tegmark M., 2000, *astro-ph/0008392*, submitted to MNRAS
- Hamilton A.J.S., Tegmark M., Padmanabhan N., 2000, MNRAS 317, L23-L27
- Haynes M.P., Giovanelli R., Chamaraux P., da Costa L.N., Freudling W., Salzer J.J., Wegner G., 1999a, AJ 117, 2039
- Haynes M.P., Giovanelli R., Salzer J.J., Wegner G., Freudling W., da Costa L.N., Herter T., Vogt N.P., 1999b, AJ 117, 1668
- Heavens A.F., Taylor A.N., 1995, MNRAS 275, 483
- Hudson M.J., Dekel A., Courteau S., Faber S.M., Willick J.A., 1995, MNRAS 274, 305
- Kashlinsky A., 1998, ApJ 492, 1
- Kolatt T., Dekel A., 1997, ApJ 479, 592
- Lin H., Kirshner R.P., Shectman S.A., Landy S.D., Oemler A., Tucker D.L., Schechter P.L., 1996, ApJ 471, 617
- Matsubara T., Szalay A.S., Landy S.D., 2000, ApJ 535, L1
- Nusser A., da Costa L.N., Branchini E., Bernardi M., Alonso M.V., Wegner G., Willmer C.N.A., Pellegrini P.S., 2000, MNRAS 320, L21
- Padmanabhan N., Tegmark M., Hamilton A.J.S., 2000, ApJ submitted *astro-ph/9911421*
- Pen U.L., 1998, ApJ 504, 601
- Primack J.R., 2000, “Cosmological parameters”, 4th International Symposium on Sources and Detection of Dark Matter in the Universe (DM 2000), Marina del Rey, California; *astro-ph/0007187*
- Santiago B.X., Strauss M.A., Lahav O., Davis M., Dressler A., Huchra J.P., 1995, ApJ 446, 457
- Saunders W., Sutherland W., Maddox S., Keeble O., Oliver S., Rowan-Robinson M., McMahon R., Efstathiou G., Tadros H., White S.D.M., Frenk C.S., Carramiñana A., Hawkins M.R.S., 2000, MNRAS 317, 55
- Sigad Y., Branchini E., Dekel A., 2000, ApJ. 540, 62
- Sigad Y., Eldar A., Dekel A., Strauss M.A., Yahil A., 1998, ApJ 495, 516
- Strauss M.A., Willick J.A., 1995, Phys. Rep. 261, 271
- Susperregi M., 2001, ApJ. 546, 85
- Sutherland W., Tadros H., Efstathiou G., Frenk C.S., Keeble O., Maddox S., McMahon R.G., Oliver S., Rowan-Robinson M., Saunders W., White S.D.M., 1999, MNRAS 308, 289
- Taruya A., 2000, ApJ 537, 37
- Taylor A.N., Watts P.I.R., “Parameter information from nonlinear cosmological fields”, *astro-ph/0010014*
- Tegmark M., 1997, Phys. Rev. Lett. 79, 3806
- Tegmark M., Bromley B., 1999, ApJ. 518, L69
- Tegmark M., Peebles P.J.E., 1998, ApJ 500, L79
- Tegmark M., Taylor A.N., Heavens A.F., 1997, ApJ 480, 22
- Tegmark M., Hamilton A.J.S., Strauss M.A., Vogeley M.S., Szalay A.S., 1998, ApJ. 499, 555
- Turner M.S., 1999, “Cosmological parameters”, in Proceedings of Particle Physics and the Universe (Cosmo-98), edited by David O. Caldwell (AIP, Woodbury, NY); *astro-ph/9904051*
- Vogeley M.S., Szalay A.S., 1996, ApJ. 465, 34
- Willick J.A., Courteau S., Faber S.M., Burstein D., Dekel A., Kolatt T., 1995, ApJ. 446, 12
- Willick J.A., Courteau S., Faber S.M., Burstein D., Dekel A., Strauss M.A., Kolatt T., 1996, ApJ. 457, 460
- Willick J.A., Courteau S., Faber S.M., Burstein D., Dekel A., Strauss M.A., 1997, ApJS 109, 333
- Zaroubi S., Zehavi I., Dekel A., Hoffman Y., Kolatt T., 1997, ApJ 486, 21
- Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ 520, 413
- Zehavi I., Dekel A., 1999, “Cosmological parameters and power spectrum from peculiar velocities”, in Proceedings of the Cosmic Flows Workshop, Victoria, B.C., Canada, eds. S. Courteau, M. Strauss, and J. Willick; *astro-ph/9909487*

APPENDIX A: LAP EQUATIONS

The cosmological perturbations are governed by (S01)

$$\mathcal{L} = \frac{1}{2}(1+\delta)\mathbf{v}^2 + \alpha \left\{ \dot{\delta} + \nabla \cdot [(1+\delta)\mathbf{v}] \right\} - \phi\delta - \frac{1}{3}\Omega_m^{-1}|\nabla\phi|^2, \quad (\text{A1})$$

where ϕ is the gravitational potential created by the matter density contrast δ . The variational equations, namely $\tilde{\delta}\mathcal{L}/\tilde{\delta}\delta = 0$ and $\tilde{\delta}\mathcal{L}/\tilde{\delta}\alpha = 0$ ($\tilde{\delta}$ denotes a variation), yield:

$$\mathbf{v} = \nabla\alpha, \quad (\text{A2})$$

$$\nabla^2\phi = \frac{3}{2}\Omega_m\delta, \quad (\text{A3})$$

and

$$\dot{\delta} + \nabla \cdot [(1+\delta)\mathbf{v}] = 0, \quad (\text{A4})$$

$$\dot{\alpha} + \frac{1}{2}|\nabla\alpha|^2 + \phi = 0. \quad (\text{A5})$$

The radius of the Universe $a(t)$ is scaled out alongside h with our use of units of km/s for the distance scale. We seek solutions for δ, α by utilizing the trial fields

$$\delta(t, \mathbf{x}) = \sum_y \delta_y(t) j_l(k_r x) Y_{lm}(\theta, \varphi), \quad (\text{A6})$$

$$\alpha(t, \mathbf{x}) = \sum_y \alpha_y(t) j_l(k_r x) Y_{lm}(\theta, \varphi), \quad (\text{A7})$$

where $y \equiv rlm$. By substituting (A6),(A7) into (A2),(A3) we get

$$\mathbf{v}(t, \mathbf{x}) = \sum_y \left[\hat{\mathbf{x}} \alpha'_y(t) j_l(k_r x) + \frac{1}{x} \hat{\mathbf{x}} \wedge \mathbf{J}_y(\alpha) \right] Y_{lm}(\theta, \varphi), \quad (\text{A8})$$

$$\phi(t, \mathbf{x}) = -\frac{3}{2}\Omega_m \sum_y k_r^{-2} \delta_y(t) j_l(k_r x) Y_{lm}(\theta, \varphi), \quad (\text{A9})$$

where the coefficients α'_y and $\mathbf{J}_y(\alpha)$ are calculated in Appendix B. The time dependence of the coefficients δ_y , α_y is modelled by a polynomial of degree N that is a sum of Chebyshev polynomials, i.e.

$$\delta_y(t) = \sum_{n=1}^N \delta_y^{(n)} T_n(t), \quad (\text{A10})$$

$$\alpha_y(t) = \sum_{n=1}^N \alpha_y^{(n)} T_n(t). \quad (\text{A11})$$

The properties of T_n are summarised in Appendix C. We do not consider the 0th order term in (A10), (A11) as it is a constant term. The polynomials (A10), (A11) that permit us to approximate δ_y , α_y , in such a way that the weighed sum of the errors in the time interval considered is least, are given by the coefficients

$$\delta_y^{(n)} = (\delta_y | T_n) (T_n | T_n)^{-1}, \quad (\text{A12})$$

$$\alpha_y^{(n)} = (\alpha_y | T_n) (T_n | T_n)^{-1}, \quad (\text{A13})$$

where $(|)$ is the discrete version of the weighed product (C13), i.e.

$$(f|g) \equiv \sum_{m=1}^M w(t_m) f(t_m) g(t_m), \quad (\text{A14})$$

where $w(t)$ is (C14) and M is the number of t_m divisions on the interval $[0, 1]$. The boundary conditions (S01)

$$0 = \sum_{n=0}^N (-1)^n \delta_y^{(n)}, \quad (\text{A15})$$

$$\rho_s(\mathbf{s}) = x^2 \left(\frac{N_{\text{gals}}}{V} \right) \left[1 + g_0(\mathbf{x}) \right] \left[1 + \alpha_0''(\mathbf{x}) \right]^{-1}. \quad (\text{A16})$$

g_0 is the present galaxy number-count density contrast and V is the volume of the survey. The coefficients δ_y enter (A16) through the bias relationship (6), that also introduces the coefficients b_{rl} . The evaluation of the modes at the present time in (A16) is simplified by virtue of $T_n(t_0) = 1$. Substituting the trial fields into equations (A4),(A5), we get

$$\sum_y (\delta_y - k_r^2 \alpha_y) j_l(k_r x) Y_{lm}(\Omega) = - \sum_{yy'} \left\{ \alpha_y' \delta_{y'} j_l(k_r x) j_{l'}(k_{r'} x) + \frac{1}{x^2} [\hat{\mathbf{x}} \wedge \mathbf{J}_y(\delta)] \cdot [\hat{\mathbf{x}} \wedge \mathbf{J}_{y'}(\alpha)] \right\} Y_{lm}(\Omega) Y_{l'm'}(\Omega), \quad (\text{A17})$$

and

$$\begin{aligned} & \sum_y \left(\frac{3}{2} \Omega_m k_r^{-2} \delta_y - \alpha_y \right) j_l(k_r x) Y_{lm}(\Omega) \\ &= \frac{1}{2} \sum_{y'} \left\{ \alpha_y' \alpha_{y'}' j_l(k_r x) j_{l'}(k_{r'} x) \right. \\ & \left. + \frac{1}{x^2} [\hat{\mathbf{x}} \wedge \mathbf{J}_y(\alpha)] \cdot [\hat{\mathbf{x}} \wedge \mathbf{J}_{y'}(\alpha)] \right\} Y_{lm}(\Omega) Y_{l'm'}(\Omega). \end{aligned} \quad (\text{A18})$$

The coefficients $\mathbf{J}_y(\delta)$ are given in Appendix B, as in $\mathbf{J}_y(\alpha)$, with the substitution $\alpha \rightarrow \delta$. By integrating out all coordinates, it is possible, though arduous, to finally arrange all the coefficients $\delta_y^{(n)}$, $\alpha_y^{(n)}$ in the following inhomogeneous matrix system for each y

$$\begin{bmatrix} \mathcal{C}_y^\alpha & \mathcal{C}_y^\delta \\ \mathcal{S}_y^\alpha & \mathcal{S}_y^\delta \end{bmatrix} \begin{bmatrix} \alpha_y^{(n)} \\ \delta_y^{(n)} \end{bmatrix} = \begin{bmatrix} \mathcal{D}_{y'y''}^y \delta_{y'}^{(m)} \alpha_{y''}^{(m')} \\ \mathcal{E}_{y'y''}^y \alpha_{y'}^{(m)} \alpha_{y''}^{(m')} \end{bmatrix}, \quad (\text{A19})$$

where \mathcal{C}_y^δ , \mathcal{C}_y^α , \mathcal{S}_y^δ , \mathcal{S}_y^α , $\mathcal{D}_{y'y''}^y$ and $\mathcal{E}_{y'y''}^y$ are computed using the orthogonality relation of the Chebyshev polynomials (C13), and the standard orthogonality relations for Y_{lm} and j_l , (D6), (D12). These terms are estimated numerically. The coefficients α_y , δ_y in the column matrix on the LHS of (A19) are arranged vertically in order of increasing n , from top to bottom, making a total of $2N$ entries. The submatrices \mathcal{C} result from the integration of the LHS of (A17), and the submatrices \mathcal{S} from integrating the LHS of (A18). The RHS of (A19) contains cross coefficients coupling different y and n . In conclusion, we have a set of constraints for $\delta_y^{(n)}$, $\alpha_y^{(n)}$: (A15), (A16) and (A19) (as we have integrated out all coordinates from the dynamical equations, (A19) is really an inhomogeneous system of constraints).

APPENDIX B: RADIAL DERIVATIVES

The radial derivative of the velocity potential can be written as

$$\frac{d}{dx} \alpha(t, \mathbf{x}) = \sum_y \alpha_y'(t) j_l(k_r x) Y_{lm}(\theta, \varphi), \quad (\text{B1})$$

where, using the equality

$$\frac{d}{du} j_l(u) = (2l+1)^{-1} [l j_{l-1}(u) - (l+1) j_{l+1}(u)], \quad (\text{B2})$$

we have

$$\alpha_y' = k_r \left[\frac{(l+1)}{(2l+3)} \alpha_{r(l+1)m} - \frac{l}{(2l-1)} \alpha_{r(l-1)m} \right]. \quad (\text{B3})$$

Similarly

$$\begin{aligned} \alpha_y'' &= k_r^2 \left\{ \frac{(l+1)}{(2l+3)} \frac{(l+2)}{(2l+5)} \alpha_{r(l+2)m} \right. \\ & \left. - \left[\frac{(l+1)^2}{(2l+3)(2l+1)} + \frac{l^2}{(2l-1)(2l+1)} \right] \alpha_y \right. \\ & \left. + \frac{l}{(2l-1)} \frac{(l-1)}{(2l-3)} \alpha_{r(l-2)m} \right\}. \end{aligned} \quad (\text{B4})$$

The coefficients $\mathbf{J}_y(\alpha)$ in (A8) are therefore given by

$$\begin{aligned} \mathbf{J}_y(\alpha) &= \frac{\alpha(l, m+1)}{2} \alpha_{rl(m+1)} (i\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2) \\ &+ \frac{\beta(l, m-1)}{2} \alpha_{rl(m-1)} (i\hat{\mathbf{x}}_1 + \hat{\mathbf{x}}_2) + im \alpha_y \hat{\mathbf{x}}_3, \end{aligned} \quad (\text{B5})$$

where

$$\alpha(l, m) = [l(l+1) - m(m-1)]^{1/2}, \quad (\text{B6})$$

$$\beta(l, m) = [l(l+1) - m(m+1)]^{1/2}. \quad (\text{B7})$$

APPENDIX C: CHEBYSHEV POLYNOMIALS

The Chebyshev polynomials of the first kind T_n satisfy the linear homogeneous second-order differential equation

$$(1-t^2) \ddot{y} - t\dot{y} + n^2 y = 0, \quad (\text{C1})$$

where the dot denotes d/dt . Following the convention of Courant & Hilbert (1989), they are explicitly given by

$$T_n(t) = \frac{n}{2} \sum_{m=0}^{[n/2]} (-1)^m \frac{(n-m-1)!}{m!(n-2m)!} (2t)^{n-2m}, \quad (\text{C2})$$

that can be rewritten as

$$T_n(\cos \theta) = \cos(n\theta). \quad (\text{C3})$$

The polynomials satisfy the orthogonality relation

$$\langle T_n | T_m \rangle \equiv \int_{-1}^1 dt w(t) T_n(t) T_m(t) = \delta_{nm} \frac{\pi}{2} (1 + \delta_{n0}), \quad (\text{C4})$$

where δ_{nm} is a Kronecker delta function and

$$w(t) = (1-t^2)^{-1/2}. \quad (\text{C5})$$

They have the following property

$$T_n(t) = T_n^* \left(\frac{1+t}{2} \right), \quad (\text{C6})$$

and their recurrence relation is given by

$$2T_n(t) T_m(t) = T_{n+m}(t) + T_{n-m}(t), \quad (\text{C7})$$

for $n \geq m$. In the asymptotic regime $t \rightarrow \infty$, the zeros of T_n are

$$t_m^{(n)} \approx \cos\left(\frac{2m-1}{2n}\pi\right), \quad (\text{C8})$$

where $t_1^{(n)} < t_2^{(n)} < \dots$ (C4)-(C8) are required in the evaluation of the numerical coefficients in (A19).

The interval of orthogonality adopted in (C4) is $-1 \leq t \leq 1$. In our application it is more convenient to scale t down to the interval $[0, 1]$, and we do so by transforming the polynomials via the change of coordinate $t = 2\tilde{t} - 1$. Therefore

$$\tilde{T}_n(t) = T_n(2t-1) = \sum_{m=0}^{[n/2]} \tilde{a}_m t^m. \quad (\text{C9})$$

The coefficients \tilde{a}_m of the new polynomials are found recursively from the old a_m through the relations

$$\tilde{a}_m^{(j)} = 2a_m^{(j-1)} - a_{m+1}^{(j)}, \quad (\text{C10})$$

for $m = n-1, n-2, \dots, j$; $j = 0, 1, 2, \dots, n$;

$$\tilde{a}_m^{(-1)} = \frac{a_m}{2}, \quad (\text{C11})$$

for $m = 0, 1, 2, \dots, n$;

$$a_m^{(j)} = 2^j a_n, \quad (\text{C12})$$

for $j = 0, 1, 2, \dots, n$; and $a_m^{(m)} = \tilde{a}_m$ for $m = 0, 1, 2, \dots, n$. Thus, the orthogonality relation (C4) in the new domain becomes

$$\langle \tilde{T}_n | \tilde{T}_m \rangle \equiv \int_0^1 dt \tilde{w}(t) \tilde{T}_n(t) \tilde{T}_m(t) = \delta_{nm} \frac{\pi}{4} (1 + \delta_{n0}), \quad (\text{C13})$$

where

$$\tilde{w}(t) = \frac{1}{2} t^{-1/2} (1-t)^{-1/2}. \quad (\text{C14})$$

Throughout our calculations we use \tilde{T}_n and the cross-product defined by (C13), as opposed to (C4), though we have omitted the tildes for simplicity.

APPENDIX D: ORTHOGONALITY RELATIONS FOR Y_{LM} AND J_L

D1 Spherical harmonics

A spherical harmonic is defined as (following the Condon-Shortley convention)

$$Y_{lm}(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!}} P_l^{|m|}(\cos\theta) e^{im\varphi} (-1)^r, \quad (\text{D1})$$

where $r = m$ for $m \geq 0$ and $r = 0$ otherwise, and the $P_l^m(x)$ are the associated Legendre functions, which are given by

$$P_l^m(x) = \frac{1}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l. \quad (\text{D2})$$

Our convention follows Courant & Hilbert (1969) and differs from Abramowitz and Stegun (1965) by a factor $(-1)^m$. (D2) can be otherwise rewritten as

$$P_l^m(x) = \frac{(-1)^m}{\Gamma(1-l)} \left(\frac{x+1}{x-1}\right)^{m/2} F(-l, l+1; 1-l; \frac{1-x}{2}), \quad (\text{D3})$$

where Γ is the gamma function and F is the hypergeometric function. The associated Legendre functions satisfy the orthogonality condition

$$\int_{-1}^1 P_l^m(x) P_n^m(x) dx = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{ln}. \quad (\text{D4})$$

The spherical harmonic satisfies the following property

$$Y_l^{-m}(\theta, \varphi) = (-1)^m Y_l^{m*}(\theta, \varphi), \quad (\text{D5})$$

and its orthonormal relation is given by

$$\int_0^{2\pi} d\varphi \int_0^\pi d(\cos\theta) Y_{lm}(\theta, \varphi) Y_{l'm'}^*(\theta, \varphi) = \delta_{ll'} \delta_{mm'}. \quad (\text{D6})$$

Furthermore, it satisfies the equality

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}(\theta, \varphi) Y_{lm}^*(\theta', \varphi'), \quad (\text{D7})$$

where the directions (θ, φ) and (θ', φ') are separated by an angle γ , and P_l are the Legendre polynomials, defined as $P_l \equiv P_l^0$.

D2 Spherical Bessel functions

The spherical Bessel function of the first kind, j_n , is defined as

$$j_l(z) = \sqrt{\frac{\pi}{2z}} J_{l+\frac{1}{2}}(z), \quad (\text{D8})$$

where $J_\nu(z)$ is the Bessel function of the first kind

$$J_l(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(l+n)!} \left(\frac{1}{2}z\right)^{l+2n}, \quad (\text{D9})$$

that is one of two linearly independent solutions of the second-order differential equation

$$z^2 w'' + z w' + (z^2 - l^2) w = 0, \quad (\text{D10})$$

where the prime denotes d/dz . One can rewrite (D8) in the so-called Gegenbauer's generalization

$$j_l(z) = \frac{1}{2} (-i)^l \int_0^\pi e^{iz \cos\theta} P_l(\cos\theta) \sin\theta d\theta. \quad (\text{D11})$$

The spherical Bessel function satisfies the following orthogonality relation

$$\int_0^1 dx x^2 j_l(k_r x) j_l(k_s x) = \frac{\pi}{4k_r} \delta_{rs} \left\{ J_{l+\frac{1}{2}}'(k_r)^2 + \left[1 - \left(\frac{l+1/2}{k_r} \right)^2 \right] J_{l+\frac{1}{2}}(k_r)^2 \right\}, \quad (\text{D12})$$

where δ_{rs} is a Kronecker delta function and $k_r \neq 0$, and also a closure relation given by

$$\int_0^{\kappa_{\max}} d\kappa \kappa^2 j_l(\kappa x) j_l(\kappa x') = \frac{\pi}{2x^2} \delta(x-x'), \quad (\text{D13})$$

where δ is in this case a Dirac delta function.

This paper has been produced using the Royal Astronomical Society/Blackwell Science L^AT_EX style file.